**1st Assignment**

**General Instructions:**

1. This is the first open book assignment.
2. You are free to consult any resource, including your friends.
3. YOU HAVE TO WRITE THE SOLUTION IN YOUR OWN WORDS. Any form of copying will result in zero marks.
4. Write your solution as a word document. You can use the equation editor to write the equations.
5. Handwritten and scanned solutions will not be entertained.
6. **Submission deadline – 5.00 p.m. on 25/9/2020.**
7. **Upload on Moodle only. Please do not send your solution by mail.**
8. Read the assignment carefully, at least twice, right up to the last full stop. Make sure that you understand the problem completely. Then start thinking about the solution.

**The Assignment:**

You have studied about the classifiers based on Bayes’ Theorem. The main steps involved are:

1. Assume that we have to build a classifier for classifying samples from K classes, C1, C2, … CK
2. Find the prior probability for each class from observed data
3. Find the class conditional probability for each class from observed data
4. Construct the posterior probability for each class using Bayes’ Theorem
5. Apply the following decision rule for classification of a test sample: Calculate the posterior probability of each class for the given sample. The class of the sample is the one that has the maximum posterior probability i.e. calculate P(C1|X), P(C2|X), … P(CK|X). Find “label” where label = argmax{P(C1|X), P(C2|X), … P(CK|X)}.

As was discussed in the class, the above rule minimizes the total misclassification error. In this assignment we are going to change this criterion. **Assume that there is a cost attached to each misclassification. Moreover, these costs are not equal.** For example, consider a two- class classifier that classifies patients as “Having Cancer – class C1” and “Not Having Cancer – class C2”. Now, if a person has cancer but our system misclassifies, the cost is very high since it can lead to the death of the patient. On the other hand, if the patient does not have cancer and our system misclassifies then the cost is smaller since the patient merely undergoes some additional tests.

Let us say you are given a cost matrix: Cost(i, j). The interpretation of the matrix elements are that if the true class is Ci and it is misclassified as Cj then we incur a cost of Cost(i, j). Note that Cost(i, j) ≠ Cost(j, i).

Assume that you are given the posterior probability for each class (i.e. steps (a) – (d) have been performed already). **In this assignment you have to derive the decision rule that minimizes the total cost of misclassification.** Now apply your derivation to the following problem.

* A classifier is to be built to distinguish between samples of two classes, C1 and C2
* Only one feature, x, is measured
* Samples of C1 are uniformly distributed in the range 0 ≤ x ≤ 4
* Samples of C2 are uniformly distributed in the range 2 ≤ x ≤ 5
* Samples of C1 are twice as likely as those of C2
* The costs of misclassification are as follows: Cost(1, 2) = 1 and Cost(2, 1) = 10
* Find the optimal decision boundary that minimizes the total error (ignoring costs)
* Find the optimal decision boundary that minimizes the total cost of misclassification